

SAM (REX/Ops) Operations Meeting 2007.12.04

Present: Adam (typist), SteveW, Randolph, Anil, Krzysztof, Angela, Parag, Robert

For reports submitted to the meeting...

The report text is in this font

Comments made during the meeting are in this font inline with the report

Releases:

Still waiting on retired date tests. Needs changes to DB.

DB server 8.4.1 is in CVS (not cut) with minor changes to request system error messages and put in "select for update" for status changes.

DØ:

Noticeable Incidents :

--- d0ora2 crashed around 9:09 am CST (Dec. 3) ... and systems seemed to be back around noon. And the Tuesday downtime was moved to Monday.

More diagnostics turned on - see DB report. Back up at 2:51pm.

--- d0ora1 host unpingable reported on Nov. 30, 2007

D0ora1 was down too on Monday. Disks were "disconnected", so system was running but not in a useful state. Back up by Monday 10:50am. Sun is investigating.

--- G. Weber (Nov. 28, 2007):

DbServers had to be restarted again after d0ora2 downtime: user_prod, web_prd_old, grid1_prd, grid2_prd, grid3_prd, grid4_prd, station_prd, user1_prd, user2_prd, user3_prd)

Some station cache that should be very fast were taking 1-2 seconds, and this slowed file delivery rates. Nelly re-ran the analysis on the tables. Not sure if this re-analysis or restarting the DB servers (or the load decreased).

On a side note - SteveW is going to ask Anil to set up cached_files and cached_file_project_usages history tables to trim the normal tables. Not sure how long it will take to move the tens of millions of records to the history table.

--- Due to problems of d0ora2 (not responding) ... the database were down between 22:04 p.m. (Nov. 27) & 14:30 (Nov. 28).

Really back by 2:51am (not 14:30).

--- OS patching has been applied to d0ora1

=====

=====

Unresolved issues:

#3210: user complained about luminosity database problem which seemed to be related to the crash on ~Monday.

Someone is continuously crashing the luminosity DB server.

Note the Lum DB server machine NFS mount their home area from d0ora2. So when d0ora2, the Lum DB server is dead too (with d0ora2 dead, can't really do much else anyway).

#3200: Large numbers of SAM Transfer failures . Maybe similar to #3199 ?

#3199 : Nov. 30: SAM was very slow (people thought it was dead) and Robert implied the database might have some abnormal behavior.

Problem solved and ticket can be closed ??

#3196 : MC OSG merge jobs cannot stage files

#3195 : samgrid jobs stalled in sam-transfer (since Nov. 28 23:36 CET).

#3188, #3181-5: Update host and service certificates for d0srv047, samgrid.fnal.gov, d0srv015, d0srv066 .

Will this be done ?

Need to renew these certs. First one expires 15 Dec and last one expires 9 Jan. If these certs are changed in production, may have some failures. Should use a group e-mail for the announcement.

SteveS: Via the API you can pass request parameters either via a file or via the python object directly. The python object method doesn't work since the SAM client will make an empty python request object if the file is not given. (See CreateRequest.params -- this method does not work; CreateRequest.paramFile works). Daniel can work around it. This problem will be fixed.

Note added in press: osg-ouhep station is in an inconsistent state. Robert is fixing and will send documentation.

CDF:

4 Inactive Cells (fcdldata320.fnal.gov) : still there

-- Nov.28 Wednesday --

1. One bad tape IAB749 in CDF-LTO3 : resolved
2. One user problem (Bizjak) : CORBA error, answered by Robert Illingworth

-- Nov.30 Friday --

1. SL8500 tape robot problem at GCC
SUN has been called to repair the failure
2. One user problem (Gorelov) : due to tape robot failure.
answered by Krzysztof.

-- Dec.01 Saturday --

1. One bad mover in LTO3 : resolved immediately
2. cdf-can station unreachable, email sent to local admin(Javier)
back up on Monday
3. Four OFFLINE cells (r-fcdfdata139-1,2,3,4) : email sent to
helpdesk, run2-sys : still there..

-- Dec.02 Sunday

1. Some upload failures in fcdfrdata321 : email sent to cdf_mc_ops
resolved soon
2. DOOR03 dcap://fcdfrdc3.fnal.gov:22128 not responding
: email sent to cdf_diskpool_admin : now resolved

Randolph: Elena has been finding duplicate luminosity in processed. Elena claims that SAM failures (Adam: Don't know about this) cause them to restart things by hand and that leads to duplicated luminosity. This is RunSection related. Problem may be that created recovery projects has subtleties and this leads to repeated data. StephanL knows about this too.

DB:

D0 :

- d0ora2 crashed on 11/26/2007 at 10:04AM. Down for 4.5 hours.
- d0ora2 crashed again on Monday 12/03/2007 ~9:10AM-2:51PM

System Log points to bad disk /backup01. /backup01 will be removed since this area is no longer used.

Sun recommended to upgrade the OBP (Open Boot Prom) which is the BIOS of the system. On top of replacing the failed drive,
Sun is recommended to install patch 119241-01.

The following is done :

Internal 73gb drive replaced because of errors
OBP upgraded to version 4.22.34
Solaris Rollup patch installed. (160 patches)
Two runs of diag set to max. No errors found.

Hence there is no need for downtime on d0ora2 Dec 04.

- OS patches on d0ora1 done on 11/27 Tuesday.
- d0ora1 was unpingable since Friday, November 30, 2007 9:49 PM. Issue was fixed Monday 12/03/06 10:15AM
Root cause is unknown. SystemAdmin is working with Sun support to find the root cause.
- Analysed the SAM schema on d0ofprd1. It took 36 hours.
- On 11/29/07 SAM cut v6_10 on d0ofdev1 - To add composite index on lower (file_name) and retired_date
 - Drop the unique index on lower(file_name)
Recreated as non-unique on file_name.
- v6_4 index on create_date of data_files against d0ofdev1 done
- Service Request with Oracle on Full table scan Vs Index scan on data_files_param_values.
Development is working on Bug Fix. Bug# 5601578 No time estimate yet for bug fix.
SAR# 5837876.993

Alerts:

Information for: d0ofprd1

Owner	Object	Type	Extents Used	Extents Allocated	Extents Left
SAMPRD	DATA_FILES_PARAM_VALUES	TABLE	3996	4000	4

Action: Increased the maxextents to 8000

CDF :

- On Nov 26, a drive failed. System Admin fixed the failed drive. No downtime.
System Admin contacted ParkPlace to replenish the cold spare.
- DB patches on cdfofdev/cdfoint done on Nov 27 8:00AM-11:00AM
- Scheduling downtime for cdfofpr2 Dec 04 8:00Am-12:00PM
DB patches fcdfora4~cdfofpr2
O/S Qrtly fcdfora4~cdfofpr2

Alerts:

MINOS:

- Scheduling the Oracle's qrtly patch on minosprd in second week of Dec/ Dec 20 (Decache may be down)

dCache/enstore:

Scheduled Downtimes:

- o D0 - Tue, Dec 4
 - no enstore work planned
- o CDF dCache - Tue, Dec 4
 - dCache will be restarted
- o SL8500 at GCC - Tue, Dec 4, 8:00 am to 10:00 am
 - an ACSLS patch will be installed to recognize the LTO4 drives

Operations:

Note added in press: rman backups have been failing since yesterday afternoon. Perhaps public dCache problem. As of 9am the rman's are working again (at least as far as being written to dCache - which had been failing before).

- o SL8500s at GCC
 - Problems with one of the handbots, culminating in its replacement on Friday. No problems since.
- o cdfdca (from Vladimir's report)
 - Door ping plots showed significant slowness of all doors starting Tuesday morning. The investigation showed that it was caused by increased numbers of file operations (~20,000/hour) on CDF PNFS server, which generated high traffic for the admin PNFS database. The situation seems to be similar what CMS had a few weeks ago.
- o fndca (from Vladimir's report)
 - There were some problems with SRM transfers on Thursday morning. Investigation showed that two SRM instances were running on fndca1 node. Both instances have been stopped and the SRM has been restarted.
 - Node stkendca9a is back after hard drive/controller repair. The dcache pools are up and running, need to be returned into PoolManager configuration.
 - 2 disks in separate raid arrays on stkendca13a went bad and were

replaced. The arrays rebuilt in both cases.